International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

Enhancing Cardiovascular Disease Prediction Based on AI and IoT Concepts

Debasish Swapnesh Kumar Nayak[1], Amrutanshu Panigrahi[2,*], Abhilash Pati[3,*], Bhimasen Moharana[4], Adyasha Rath[5], Ghanashyam Sahoo[6], Manoranjan Dash[7]

[1,2,3,6] Dept. of CSE, FET-ITER, Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India
[4] School of CSE, Lovely Professional University, Punjab, India
[5] Department of CSE, C. V. Raman Global University, Bhubaneswar, India
[6] Department of CSE, GITA Autonomous College, Bhubaneswar, India
[7] Faculty of Management Sciences, Siksha 'O' Anusandhan University, Bhubaneswar, India
* Corresponding author's Email: er.abhilash.pati@gmail.com, amrutansup89@gmail.com

Abstract— One-third of all deaths worldwide yearly are attributable to cardiovascular disease (CVD). In contrast to the 7% of the wealthy who experience premature death, 43% of the poor do. Lifestyle diseases like obesity and diabetes are to blame. The importance of early identification of heart disease was demonstrated, and premature mortality was kept to a minimum. Combining clinical and biochemical data is essential for the early diagnosis of heart illness. Numerous IoT-enabled wearable healthcare applications have been created and released in recent years. Although the ability of wearable devices to share patient health data is expanding, it remains challenging to predict and identify health problems. Security, data storage, and patient monitoring are all part of the system. Artificial intelligence (AI) therapies may one day change the face of cardiology by providing doctors with cutting-edge data analysis and therapeutic decision-making resources. As the volume and complexity of data continue to increase, AI tools like machine learning (ML) and deep learning (DL) can assist medical professionals in learning more. Suppose we want to provide medical care to the elderly and those with chronic illnesses in the comfort of their own homes. In that case, we must upgrade our communication and information technology systems. The implemented DNN model's accuracy is amazing at 95.34 % and can yield other noteworthy outcomes when used to identify CVDs. We discuss and suggest the most suitable AI-IoT models for early CVD prediction and detection to reduce computational costs and increase time efficiency.

Keywords- Deep Learning (DL); Machine Learning (ML); Internet of Things (IoT); Artificial Intelligence (AI); cardiovascular disease (CVD)

I. INTRODUCTION

Health is characterized not by the absence of disease but by the presence of functional physical and cognitive systems. The modern world's death rate has increased due to the rise of non-communicable diseases. Cardiovascular illness takes first place worldwide among chronic, non-communicable diseases. According to data compiled by the American Heart Association (AHA), cardiovascular disease (CVD) is responsible for one in four fatalities in the United States. According to the World Health Organisation (WHO), heart disease kills 12 million people annually, or around 40% of the global population. If fatalities could be anticipated, many lives may be saved. Therefore, quick medical intervention is crucial. Designs for healthcare systems are needed to effect change [1]. CVD is a problem for both men and women beyond 35. Alcohol, tobacco use, obesity, hypertension, diabetes, and high cholesterol all contribute to one another. Recent research confirms this to be the sole cause of cardiac arrest. Heart failure is the most common cause of death. The heart's ability to pump blood to the ventricles and empty them is impaired. If the heart stops beating due to a lack of blood, it damages the eyes, kidneys, and brain and can be fatal [2, 3]. The Internet of Things (IoT) is a game-changing platform supported by multiple fields of study that generates premium connectivity between many interdependent things to better the user experience. By bringing together all of the relevant parties and employing cutting-edge technology to maximize the value of data shared among tightly networked nodes on the IoT platform, the healthcare system can be rethought from the ground up. The IoT is crucial in healthcare since it helps reduce stress for both patients and doctors. It comprises a system that integrates networked systems, apps, and devices to assist patients and physicians in monitoring, tracking, and documenting vital data and medical information. Some examples of these devices include sensors built into smartphones and Wi-Fi dongles. Smartphone apps provide timely alerts and emergency support and help keep medical records. The large amounts of data produced by these interconnected IoT devices provide a serious challenge to the operator. The ability to recognize when someone is in peril is crucial. Clinical diagnosis is a crucial process that requires precision and prompt action [4, 5]. Automation is essential to help doctors make more accurate diagnoses because human intelligence can make false assumptions and yield unpredictable results. It's a tool for

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

doctors, and it helps patients know when they should seek medical attention. Therefore, illness analysis plays a significant role in healthcare [6, 7]. Patient records and other databases may contain a hospital's wealth of knowledge. There were 271 million CVD diagnoses in 1990; there are 523 million in 2019, and CVD-related fatalities have gradually increased from 12.1 million to 18.6 million. Due to changing demographics for DALYs and years of life lost, years lived with disability increased from 17.7 million to 34.4 million. DALYs from IHD have climbed steadily since 1990, with 182 million in 2019, 9.14 fatalities, and 197 million regular IHD cases. Strokes caused around 143 million DALYs in 2019. This includes 101 million frequent strokes and 6.55 million deaths annually. Fig. 1 depicts condensed CVD instances for the world's top seven countries. Early detection, diagnosis, and treatment of CVD are all made possible by artificial intelligence (AI) technologies, including machine learning (ML), deep learning (DL), etc. Large datasets (numerical, qualitative, and transactional data) have resulted from massive EHR data purchases.

A. AI and IoT in Smart Healthcare

Data prediction, detection, analysis, and classification challenges are met using an IoT strategy. The ability to use IoT over a long period of time has made treating a large number of patients quickly and easily. Indicators of physical status may be measured, monitored, and analyzed with precision thanks to sensors that harness the power of the IoT. Twenty years ago, the first IoT solution in healthcare emerged, and its potential to enhance service quality while decreasing patient costs was immediately apparent. Gartner Analytics has identified IoT in healthcare as a key technological trend for 2020. By 2025, the healthcare IoT market might be worth $534.3 billion. In addition, as can be shown in Fig. 2, 60% of healthcare providers are rapidly deploying IoT solutions.

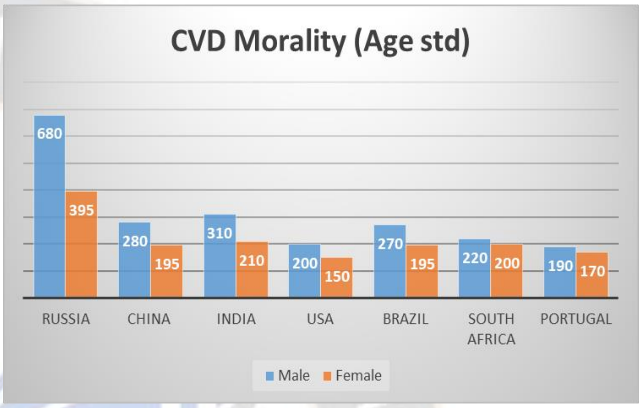

Figure 1. CVD cases for the top seven countries in 2021 (Male: Female)

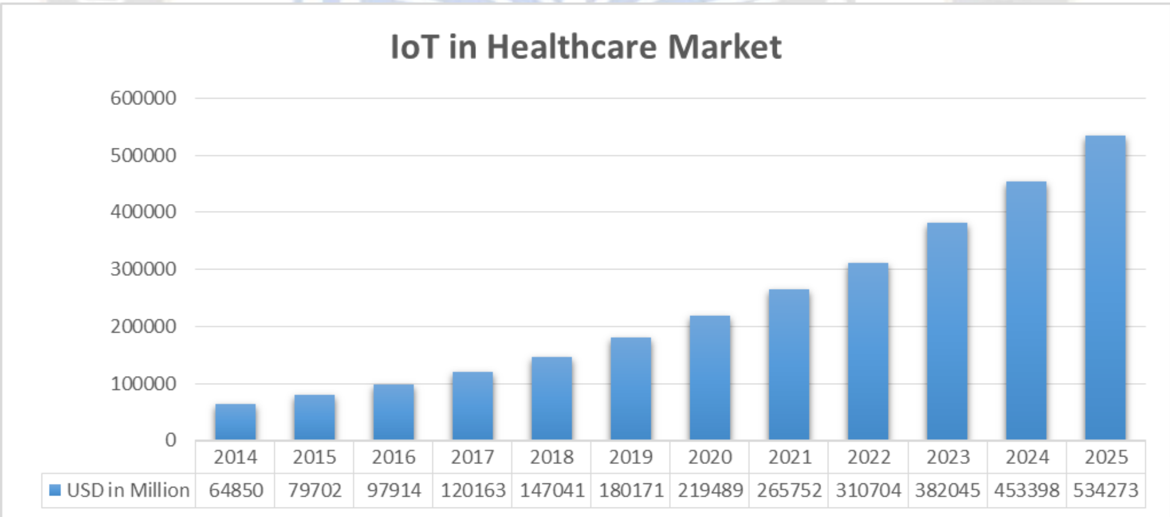

Figure 2. Prediction of Global IoT Market 2025

New computational intelligence methods can be put into practice more quickly and with less difficulty, thanks to the ever-improving state of computing technology. Better healthcare diagnoses can be made with the help of AI, ML, and DL techniques applied to data collected from IoT sources. Biomedical data can be complex and large, so it's vital to examine many models to find the ones that work best with the information. When applied to biological data, ML methods such as Decision tree (DT), random forest (RF), Support vector machine (SVM), and Gradient boosting (GB) have shown promising results. Artificial neural networks (ANNs), backpropagation neural networks (BNNs), and deep neural networks (DNNs) are all examples of DL methods that have shown promise for analyzing medical data, particularly data gathered from sensors. As shown in Fig. 3, the market value for AI in healthcare is expected to grow to a whopping USD 74,650.7 million by 2027. Major rivals in the field of AI in the healthcare business worldwide include IBM Corp., Medtronic

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

Plc., Google LLC, Koninklijke Philips N.V., Microsoft Corp., Nvidia Corp., iCarbonX, Atomwise Inc., and CloudMedx Inc.

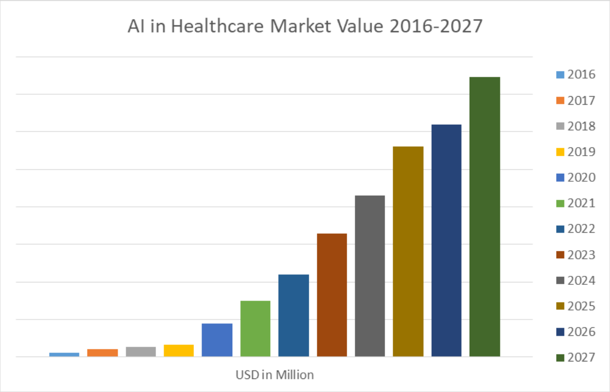

Figure 3. AI in the Healthcare Market

B. *Paper Structure*

The following is how the paper is organized: The second section compares and contrasts pertinent works to this study. The third section includes the dataset, data pre-processing techniques, and algorithm. In the fourth section, model architecture and implementation are covered. The final portion contains a conclusion with the future plan.

II. LITERATURE SURVEY

Chatterjee et al. [8] studied IoT and intelligent healthcare decision support systems. Work on an IoT healthcare platform and decision support system that satisfies the goals of an efficient technology-led healthcare system and identify risk groupings from a sample of people. People were grouped into risk groups for CVD using this method. The risk classifications were based on the likelihood of heart disease. This technology can be extended to additional sectors and give a common platform for all healthcare stakeholders. Jabeen et al. [9] developed an IoT-based community-based recommender system that diagnoses heart disease and provides physical and dietary counseling. The first phase collects data from the patient using biosensors. IoT sends data to the server. On the other hand, if you have a heart illness, you can use a heart disease prediction model to identify and classify it into eight categories (SVT). It also provides physical and dietary counseling for cardiac patients based on gender and age. A renowned hospital's specialized cardiologist compiles an illness database. Precisely 98 percent in precision, recall, and mean absolute error. Gia et al. [10] showed a fog-based remote health monitoring and fall detection system. The device can also remotely monitor room temperature, humidity, and air quality. ECG features, security, and locally distributed storage are extracted via Fog computing at the network edge. A precise and energy-efficient sensor node. Despite its many sensors, it can run for 157 hours on a 1000 mAh lithium battery. Patro et al. [11] provide a system for predicting cardiac disease utilizing K-nearest neighbors, Naive Bayes, SVM, Lasso, and ridge regression methods. Analyses of linear discriminant and principal components The SVM is 92% accurate, and the F1 is 85%. The suggested study's precision, accuracy, and sensitivity are evaluated. Ahmid and Kazar [12] devised a cloud-IoT system based on agents for remote heart rate monitoring to track and monitor patients with cardiovascular illnesses anywhere. We present a system to record and analyze patient heart rates. For emergencies, it can make speedy decisions. It is not limited to hospitals. Huang et al. [13] used wearable sensors to collect daily patient data. The remote doctor will order cardiovascular imaging if the monitoring data is abnormal. Edge computing is used to classify these training photographs and generate pseudo-labels. Stent struts are also segmented in intravascular OCT and intravenous ultrasonography images of patients. To complete telesurgery, remote and local clinicians visually communicate. We employed a U-net backbone with a SeResNet34 encoder to segment the stent struts. Meanwhile, we tested our method's accuracy, sensitivity, Jaccard, and dice. Wang et al. [14] developed a new wireless ECG patch using CNN and LSTM models. We discovered that present algorithms cannot distinguish two primary heartbeat types (Supraventricular premature beat and Atrial fibrillation) with 58.0% accuracy. Unlabeled data can be trained semi-supervised using confidence levels. The suggested method is 90.2 percent accurate on average and 5.4 percent more accurate than current ECG categorization procedures.

III. MATERIALS AND METHODS

The website's valuable data can be obtained and trained to anticipate the patient's health. The parameters have been adjusted to the model, and the outcome is appropriate. As a result, the test data may be used to evaluate the model's performance. This section explains how prognosis data is obtained, pre-processed, and subsequently trained.

A. *Data Collection*

The proposed system is trained with the integrated dataset of four heart disease datasets (HDDs), namely, Hungarian, Cleveland, Switzerland, LongBeach, and Starlog, obtained from the UCI-ML repository [15]. A short description of the datasets used is depicted in Table 1.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

TABLE I. CHARACTERISTICS OF THE CVD DATASETS

Dataset Name	Attributes	Instances	Healthy	Illness	Class Label
Hungarian	14	294	188	106	2
Cleveland	14	303	164	139	2
Switzerland	14	123	08	115	2
LongBeach	14	200	51	149	2
Starlog	14	270	150	120	2
Integrated CVD (Pre-processed)	12	1190	561	629	2

There were 294 occurrences in the Hungarian dataset, with 188 instances belonging to the healthy group and 106 to the heart disease group. In the Cleveland dataset, there were 303 cases, with 164 instances belonging to the healthy and 139 to the heart illness. There were 123 cases in the Switzerland dataset, with 08 cases belonging to the healthy and 115 cases belonging to the cardiac ailment. There were 200 cases in the LongBeach dataset, with 51 occurrences belonging to the healthy and 149 cases of cardiac disease. The Starlog dataset has 270 cases, of which 150 are healthy and 120 are diseased. For each case, 14 clinical characteristics were collected, as shown in Table 2. The final dataset is prepared after data processing, and it is the integration of all of these five datasets. The integrated dataset has a total of 1190 cases, out of which 561 are healthy and 629 are illness cases; there is a total of 12 clinical characteristics in this dataset, and it has 02 class labels, as shown in Table 1.

B. Data Pre-processing

Raw data is transformed into a machine-readable format using data preparation techniques. The data obtained about patients in the actual world is frequently inconsistent, incomplete, and likely to contain numerous inaccuracies. Data preparation approaches can help with these challenges. The strategies for data pre-treatment that are employed in phase-I and phase-II of the model implementation here are.

- Imputation (Phase-I)
- Normalization (Phase-I)
- Feature selection (Phase-I)
- Encoding in Real-Time (Phase II)

1) Imputation

The data in this set is ambiguous and unreliable. In the dataset, a few attributes may contain sensitive patient details, which must be removed because they aid the neural network's generalization. That specific tuple can be handled if the class label is absent. When a value other than the target (class label) is missing, that value can be manually filled in. Attribute means can also be used to fill in missing values. Manual imputation is used on the datasets because attribute means imputation of missing values is time-consuming.

2) Normalization

In many circumstances, normalization is advantageous. It enhances the model's numerical stability while also minimizing training time. When numbers in a bounded interval are required, normalization using the min-max scaling approach is commonly used. Since many linear methods have weights set to 0 or very few random values close to 0, ML methods are more amenable to standardization. Standardization places a mean of zero and a standard deviation of one on the feature values, creating a normal distribution across the feature columns. In addition, standardization preserves important information about outliers and makes the technique less sensitive than min-max scaling, which restricts data to a range of values. In scikit-learn, standardization is implemented via the StandardScaler class, which is analogous to MinMaxScaler. Also, the StandardScaler parameters can only be used to transform the training data, not the test set or any additional data point [28].

3) Feature Selection

Identifying the most significant or similar qualities is known as feature selection. A disease prediction cannot be made using the attributes given in a cardiac dataset. It also necessitates more storage and more time for training and testing the data, and its correctness varies. A model's accuracy can be improved by reducing the number of attributes. Only a few characteristics are required to predict CVD, and the process is substantially faster. There are numerous strategies for identifying important attributes, but data augmentation can be used because it improves feature selection accuracy. Supervised learning is an entropy-based feature selection strategy that eliminates the dataset's irrelevant features. It also minimizes the amount of noise produced.

4) Encoding in Real-Time

Data is translated into an acceptable format for analysis in this step. Using the hierarchy's paradigm, higher-level data replaces the lower-level data. The integer encoded field is discarded for each integer value, and a new binary constant of 1 or 0 is added. The attribute data is scaled from the dataset to fall between 0 and 1. The age of lower-level data, for example, can be classified into three categories: (0) youth, (1) middle-aged, and (2) senior. The metadata smoker might be mapped between

0 (no smoking) and 1 (smoking) (when an individual smokes). This step is performed in implementing the proposed model's second phase.

C. AI for CVD Detection and Prediction

AI is a computer system that can do tasks based on data it receives from other sources by employing techniques like algorithms, cognitive computing, heuristics, pattern matching, rules, and DL. The term AI was created in 1956 by scientists at Dartmouth College. As a "brain model" for the supervised training of binary classifiers, Rosenblatt developed the perceptron in 1958. With its groundbreaking backpropagation learning method, neuron-like units may be networked to learn any function. Using a deep convolutional neural network for object recognition and GPUs (graphics processing units) to accelerate network training, Crizhevsky and colleagues won the ImageNet ILSVRC Challenge 2012 in 2012. In the field of cardiovascular disease medicine, artificial intelligence research focuses on imaging. ML models are useful in echocardiography because they help standardize results between operators and reveal unique information that is hard for humans to detect. AI has the potential to be used in cardiac CT for those who are suspected of having CAD. Clinicians can learn more about their patients' functionality beyond just the atherosclerotic plaque characterization using cardiovascular CT and ML algorithms. Electrocardiogram (ECG) abnormalities may potentially be detected using ML in cardiology. Experts are interested in developing fully integrated care due to AI's rapid development, especially in ML and DL modules. The healthcare provided by these dependable and more effective methods is excellent. As a result of their mutually reinforcing definitions, ML and DL have given rise to approaches that are finding ever-growing applications in the field of cognitive computing. Before ML, applications simply followed algorithmic instructions, and implementations were associated with particular code. However, if additional rules were to be added, the code would have to be largely rebuilt. ML allows for more comprehensive techniques, as a model may be created that offers an intelligent and accurate solution to learn new rules. The basic algorithms employed in ML are mostly statistical. DL is a subsection of ML that deals with the application of neural networks, with the deep term referring to the count of layers in the network [16, 17].

1) Machine Learning (ML)

ML is a subset of AI that intends to "train" computers to rapidly, correctly, and efficiently evaluate enormous datasets using advanced computational and statistical methodologies. These algorithms typically recognize patterns in new data that match patterns in previously "learned" evidence and create predictions based on the procedure [18, 19]. The predictive model is a functional relationship between input (x) and output (y) in ML with a functional relation y = F(x). The ML is broadly classified into three groups depending on the prediction model's data accumulation and learning type [19, 20].

I. Supervised learning (e.g., SVM, LR, and NN) uses human-labeled values (datasets) to create models to predict or classify future events or identify the most important variables in determining the results. Both input (x) and output (y) are known, and data training enhances and benefits the predictive model.

II. Unsupervised learning, the technique can uncover latent structures in the datasets without categorizing the training set previously (only input x is known). This is the ability of unsupervised learning to uncover new connections in the data.

III. Reinforcement learning is a sort of reward-based learning in which true and false reinforcements are used to improve the predictive model. It is commonly employed in gaming and robotics. It necessitates the inclusion of technologies and tools that allow the machine to increase its learning and comprehend its surroundings' features, such as cameras, sensors, and GPS.

2) Deep Learning (DL)

DL is a supervised ML technique that employs neural networks to find significant patterns in large data sets [21, 22]. Like the human brain, it can construct sophisticated feature representations from data. The developer inputs new data into the model so that designed algorithms work properly even when faced with new information. An upgraded input-output level neural network learns by doing, reading data, and building hierarchical topologies. It can handle nonlinear input-output interactions. This way, the mean error of outcomes and forecasts can be minimized.

IV. PROPOSED AI-IOT HEALTHCARE MODEL

The proposed model has two different phases. At first, the raw data was fed to the model in the ratio of 80:20 (train: test). In the later phase of model evaluation, real-time data was taken to identify the accuracy of the implemented model in identifying CVDs.

A. Phase-I

The proposed model has two different phases; in the first phase, the developed model is trained and tested on the data gathered from the open-access repository in an 80:20 ratio. Different parameters of the prediction model are being measured, as shown in Fig. 4.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

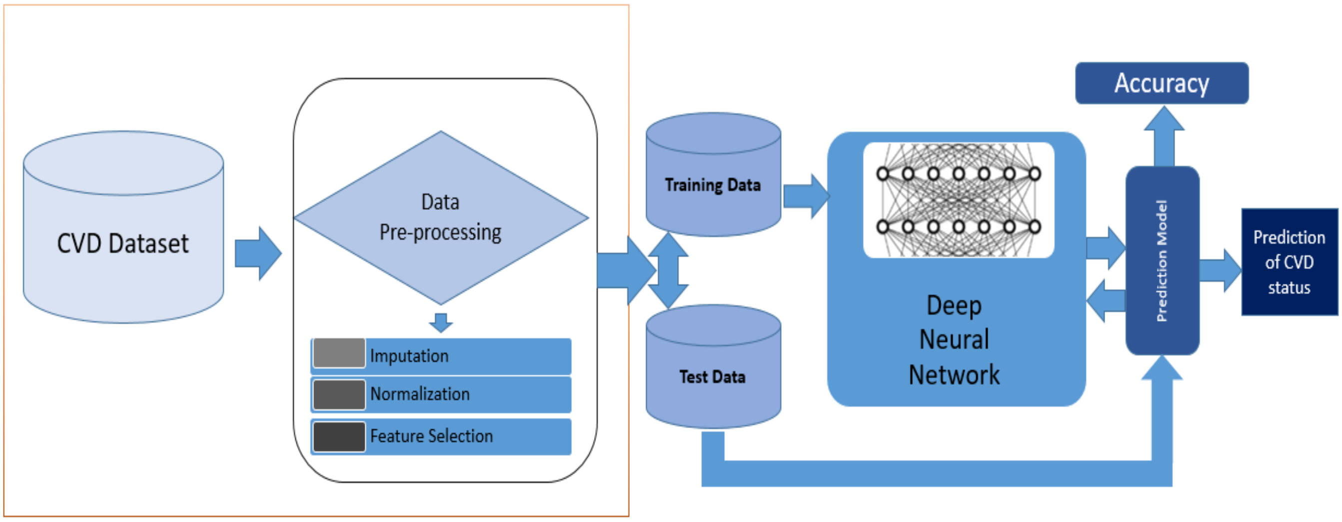

Figure 4. Blueprint of the Proposed Model (Phase-I)

Implementation of Artificial Neural Network (ANN) and Deep Neural Network (DNN)

Anaconda navigation is installed and utilized because it comes with built-in Python packages, is easy to use (user-friendly), and is simple to start with (set up). It contains a Jupyter notebook and a spyder. Spyder with Tensorflow is utilized here, and the ANN application is written in Python. Spyder makes debugging considerably smoother [23, 24]. Hence, we implement the model with a spyder environment. ANNs have been widely used in CVD prediction and risk assessment. ANNs are a type of machine learning model that can capture complex patterns in data, making them suitable for tasks like CVD prediction. A DNN that solves non-linear problems is desirable. It belongs to the ANNs category. As previously stated, the method has an input, hidden, and output layer. The number of nodes in hidden layers may be any number but should not exceed the number of input neurons. The accuracy can be adjusted by adjusting (increasing/decreasing) the number of nodes in the hidden layer. The learning rate is 0.5, the input layer has 12 neurons, and there are 50 training epochs. Because a system is more complex with a higher learning rate, it should be kept to a minimum. The mean square error (MSE) is calculated for each iteration. The MSE acquired at the end of epochs is 20.31, and the cost function is 1.39. The mean square error value decreases with each iteration. Furthermore, as the number of iterations is raised, accuracy increases as well, i.e., the MSE value is inversely related to the number of iterations, but the accuracy is directly proportional.

B. Phase-II

In the second phase, the prediction model is fed with the real-time data collected from various sensing devices (remote places) and personal patient records available in the hospital, as shown in Fig. 5.

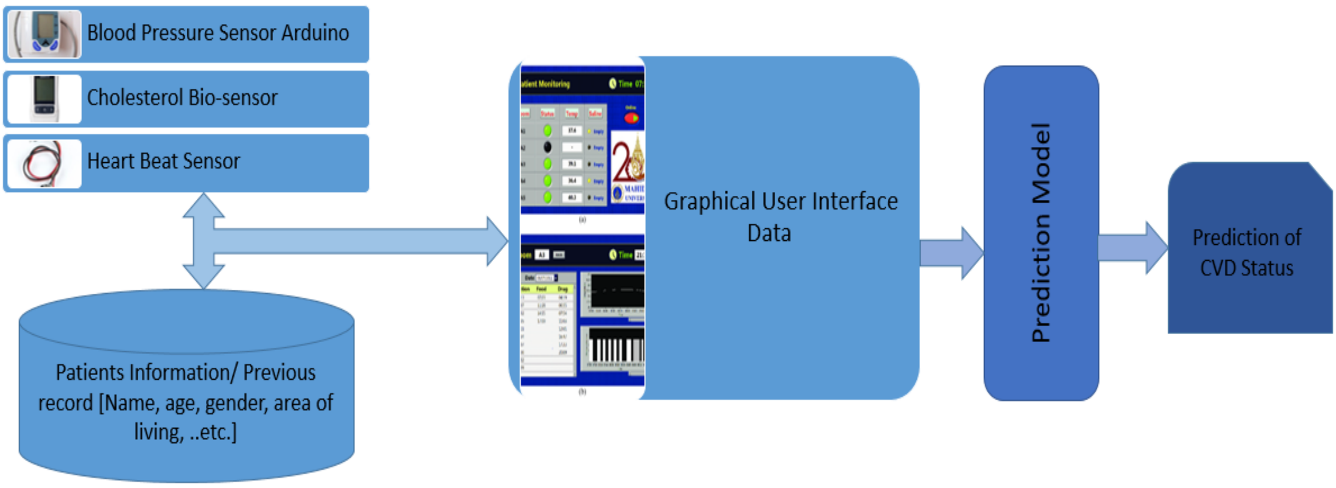

Figure 5. Blueprint of the Proposed Model (Phase II)

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

The real-time data was collected using Arduino. In the actual and virtual worlds, Arduino creates single-board microcontroller kits to make digital devices and interactive items to watch and control objects. An association for the advancement (shields) and other circuits are connected to the board via digital and analog input/output (I/O). The board includes 14 digital and six analog pins and can be programmed using the Arduino IDE via a USB type B connector. It may be charged via a USB port or a 9-volt auxiliary battery.

V. RESULTS AND DISCUSSION

The results obtained using this proposed approach and a discussion along with the exploratory data analysis are discussed in this section. The proposed model is developed on a workstation with 32GB of RAM, 1 TB of SSD storage, an Intel Core i7 processor, and a Ubuntu 20.04 operating system. The predictive model's performance can be evaluated in several ways [25-27]. This study was conducted with the Integrated CVD dataset in mind, along with four other HDDs that can be downloaded for free from the UCI-ML repository and a handful of performance metrics like accuracy, precision, sensitivity, specificity, F1-Score, etc., as shown in Table 2. Using the Confusion matrix (CM), as depicted in Table 3, acquired from the Jupiter Python Notebook experiments [28-31], we can calculate these performance parameters, as shown in equations (1)-(12).

TABLE II. EVALUATIVE PARAMETERS AND USED NOTATIONS IN TRAINING THE PROPOSED MODEL

Parameters	Notation Used
Accuracy	Ac_y
Misclassification	M_r
Precision	P_n
Sensitivity	S_{ey}
Specificity	S_{py}
F1-Score	$F1_s$
False Negative Rate (FNR)	F_N
False Positive Rate (FPR)	F_P
Mathews Correlation Coefficient (MCC)	M_C
Kappa	κ
Relative Absolute Error (RAE)	RA_E
Root Relative Squared Error (RRSE)	RRS_E

TABLE III. CONFUSION MATRIX CHARACTERIZATION IN THE STUDY

		Observed Value	
Predicted Values		True	False
	True	T	F
	False	$\tilde{F}$	$\tilde{T}$

$$Ac_y = \left(\frac{T+F}{T+\tilde{T}+F+\tilde{F}}\right) \quad (1)$$

$$M_r = \left(1 - Ac_y\right) \quad (2)$$

$$P_n = \frac{T}{T+F} \quad (3)$$

$$S_{ey} = \frac{T}{T+\tilde{F}} \quad (4)$$

$$S_{py} = \frac{\tilde{T}}{\tilde{T}+F} \quad (5)$$

$$F1_s = \frac{T}{T+\left(\frac{F+\tilde{F}}{2}\right)} \quad (6)$$

$$F_N = \frac{F}{T+F} \quad (7)$$

$$F_P = \frac{\tilde{F}}{\tilde{T}+\tilde{F}} \quad (8)$$

$$M_C = \frac{\{(T*\tilde{T})-(F*\tilde{F})\}}{\sqrt{(T+F)(T+\tilde{F})(\tilde{T}+F)(\tilde{T}+\tilde{F})}} \quad (9)$$

$$\kappa = \frac{2*\{(T*\tilde{T})-(F*\tilde{F})\}}{(T+F)(T+\tilde{F})(\tilde{T}+F)(\tilde{T}+\tilde{F})} \quad (10)$$

$$RA_E = \frac{\sum_{i=1}^{n}\left|p_{f_i}-a_{f_i}\right|}{\sum_{i=1}^{n}\left|\bar{A}-a_{f_i}\right|} \quad (11)$$

$$RRS_E = \sqrt{\frac{\sum_{i=1}^{n}\left|a_{f_i}-p_{f_i}\right|}{\sum_{i=1}^{n}\left|a_{f_i}-\bar{A}\right|}} \quad (12)$$

$$\bar{A} = \frac{\sum_{i=1}^{n} a_{f_i}}{n} \quad (13)$$

Where $p_{f_i}$is the predicted value of feature f_i, a_{f_i} is actual value of feature f_i, and $\bar{A}$ is the mean of actual values.

A. Exploratory Data Analysis (EDA)

The heat map visualizes the relationship between the various feature values in the integrated CVD dataset, as shown in Fig. 6. The highly correlated and less correlated features are marked with red and black, respectively. The correlation among various cases is analyzed. The clinical feature “resting bp” is taken as the base feature and compared with four major features of the dataset: “cholesterol,” “heart rate,” “chest pain type,” and “ST slope,” taken with class labels as the correlation factor. It is observed that the correlation is very good among the “cholesterol” - “chest pain type” and “resting bp” - “resting bp.” At the same time, the correlation factor in the other two cases is less compared with the first two cases, as shown in Fig. 7. Fig. 8 shows the bar plot of the class label target attributes.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

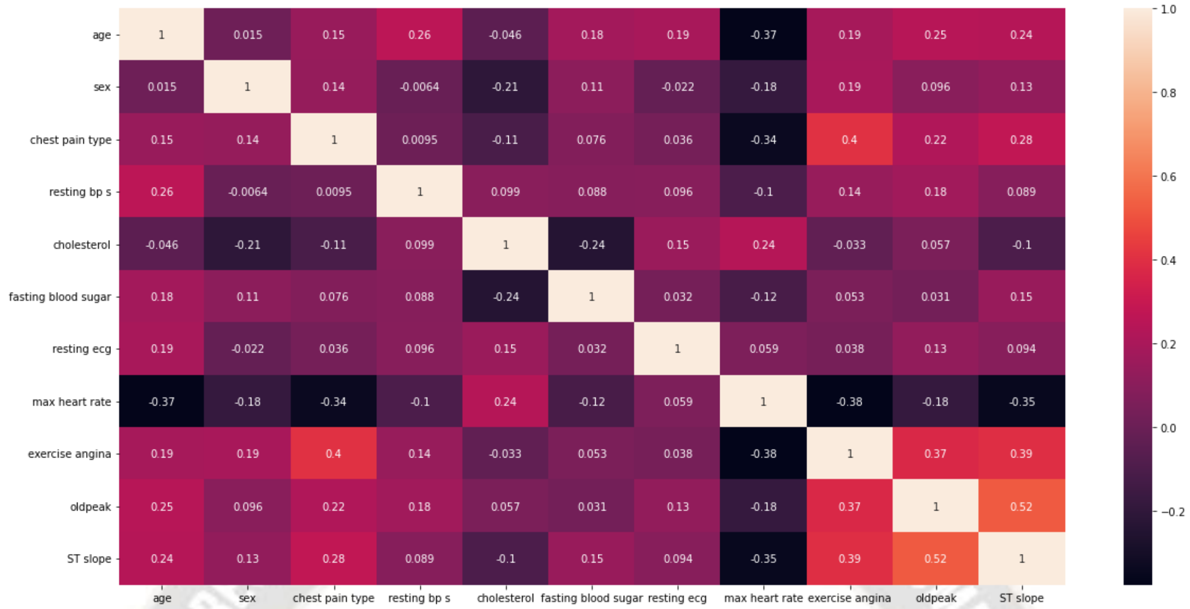

Figure 6. Heatmap of the Pre-processed Integrated CVD Dataset

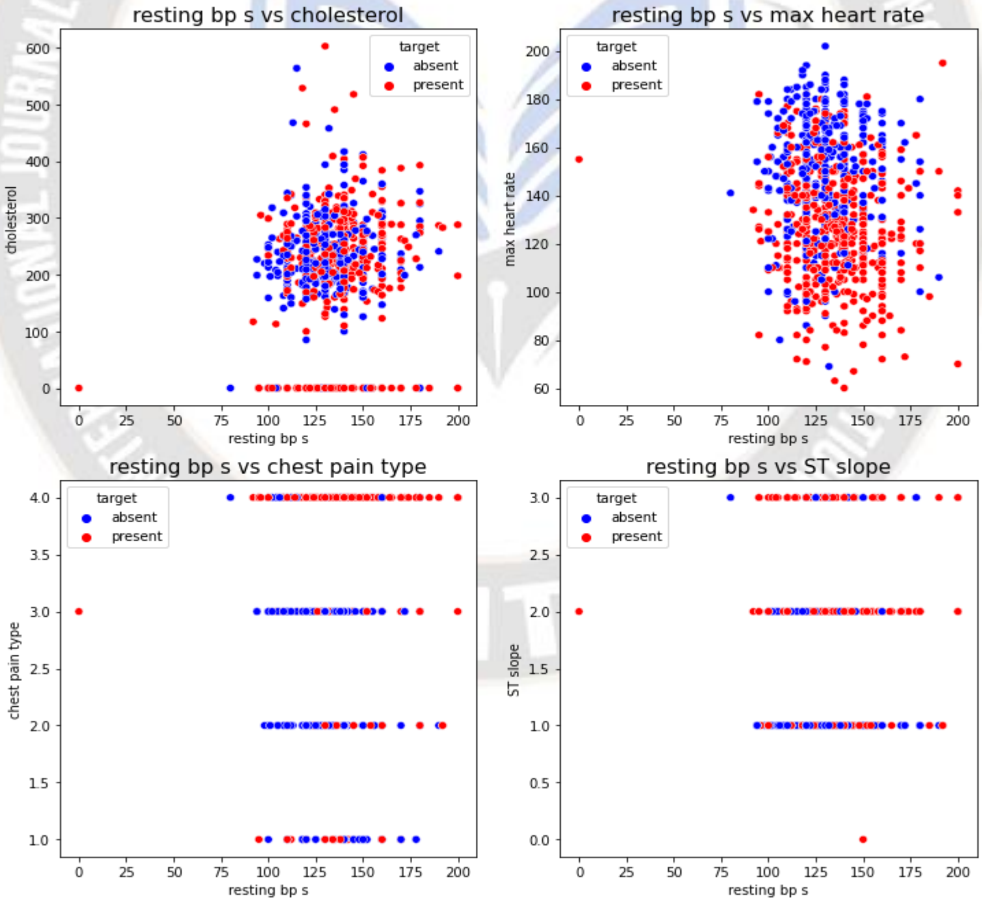

Figure 7. Scatter plot for correlation among different features taking class label as the factor of correlation

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

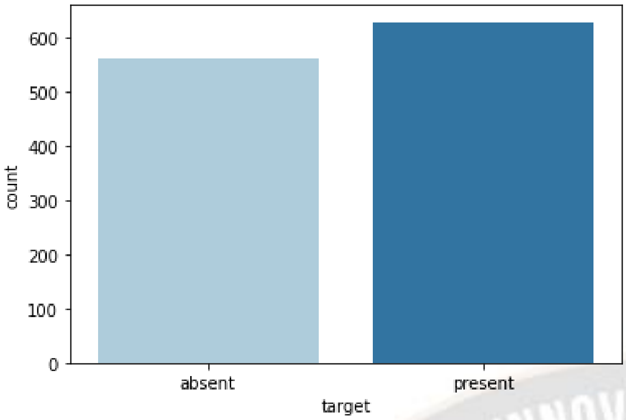

Figure 8. Bar plot of the target attributes (class label)

B. *Analysis of Obtained Results*

The integrated CVD dataset is divided into an 8:2 ratio. The training data is fed into an ANN that employs the backpropagation method. The DNN algorithm is used to test the data. Three activation functions, ReLu, Sigmoid, and Tanh, are used to test an algorithm's accuracy, and the optimal activation function, ReLu, is picked for the model as it gives the highest accuracy. Dealing with complex datasets, the DL models mostly use two commonly used activation functions, sigmoid and tanh. But, saturation is a major issue, mostly with the sigmoid and tanh functions. It becomes harder for the learning algorithm to alter the weights to increase the model's performance once it becomes saturated. Whereas ReLu allows for increased sensitivities to the activated sum input while avoiding saturation and thus helps to increase the model accuracy. The optimizer used for the model is Adam (Adaptive Moment Estimation Algorithm); this algorithm estimates moments and optimizes a function using them. It combines the gradient descent with the momentum technique and the RMS (Root Mean Square) method. The Adam technique squares the obtained gradient after calculating an exponentially weighted moving average. Two decay factors in this technique regulate the drop rates of these created moving averages. The Adam optimizer best fits the model and produces high accuracy in training and test cases. Different evaluating parameters of the implemented model. Table 4 and Figures 9 and 10 summarize the results obtained using this suggested approach. The important factor with the implemented DNN model is the accuracy, which is a remarkable 95.34% and can produce other significant results for identifying CVDs.

TABLE IV. DETAILS OF RESULTS OBTAINED FROM THE IMPLEMENTED ANN AND DNN MODELS

Evaluative Parameters	Results using ANN (in %)	Results using DNN (in %)
Accuracy	89.75	95.34
Misclassification	10.25	4.66
Precision	91.36	95.28
Sensitivity	93.49	98.23
Specificity	82.24	88.54
F1-Score	92.41	96.73
FNR	6.51	1.77
FPR	17.76	11.46
MCC	76.68	88.76
Kappa	73.15	78.19
RAE	44.11	49.47
RRSE	63.56	68.56

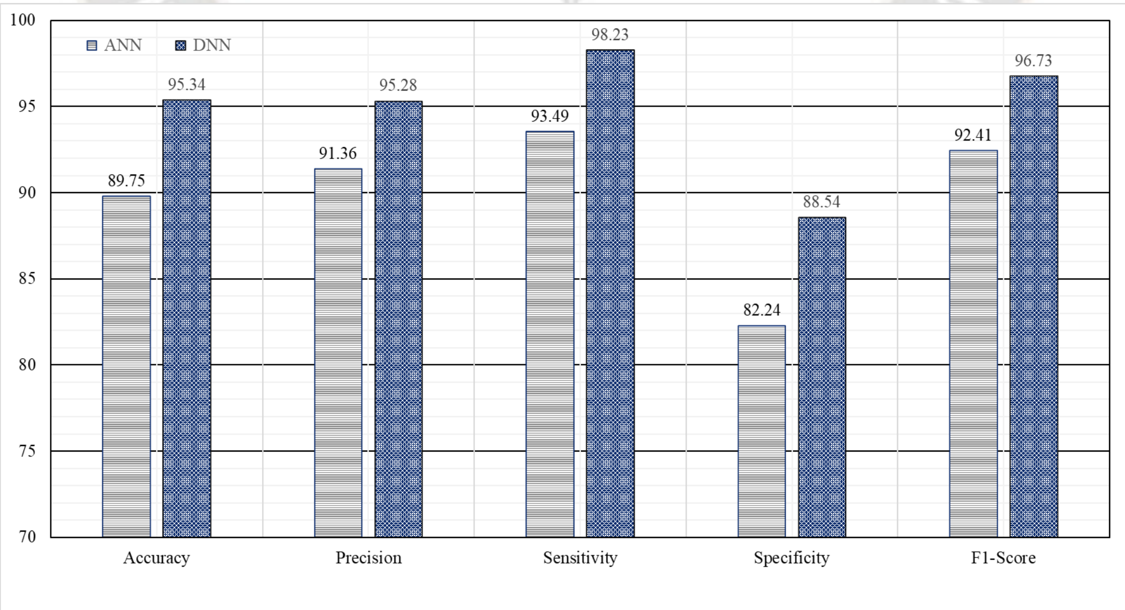

Figure 9. Comparative Analysis of Results Obtained using the ANN and DNN Model (in %) in terms of Accuracy, Precision, Sensitivity, Specificity, and F1-Scores

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

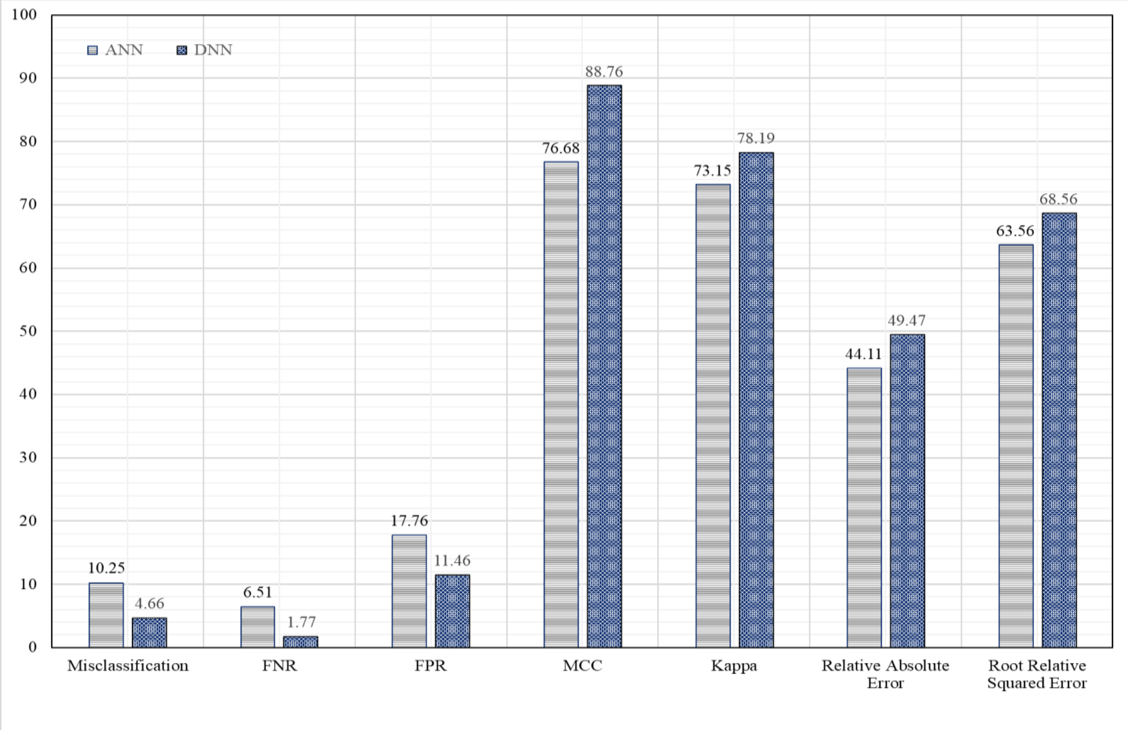

Figure 10. Comparative Analysis of Results Obtained using the ANN and DNN Model (in %) in terms of MCR, FNR, FPR, MCC, Kappa, RAE, and RRSE.

VI. CONCLUSION AND FUTURE SCOPE

Nowadays, cardiac arrest requires special attention since statistics show that the number of deaths caused by CVD has climbed significantly in recent years. Early heart attack prediction and preventative actions are required to prevent earlier deaths. DL approaches have demonstrated their ability to detect CVDs early and accurately. The accuracy of the implemented DNN model, at a remarkable 95.34 %, is the most crucial factor, and it can provide other notable outcomes, as from the outcomes obtained, for diagnosing CVDs. The model's most important feature is that it does not require any medical features gathered through various medical examinations (which may increase the cost). In the future, we intend to construct a hybrid model that retains and improves the prediction algorithm employed in this study to attain improved prediction accuracy.

REFERENCES

[1] M. S. Atoum, A. Pati, M. Parhi, B. K. Pattanayak, A. K. Habboush, M. Alnabhan, E. Qalaja, "A Fog-Enabled Framework for Ensemble Machine Learning-Based Real-Time Heart Patient Diagnosis," International Journal of Engineering Trends and Technology, vol. 71, no. 8, pp. 39-47, 2023. doi: https://doi.org/10.14445/22315381/IJETT-V71I8P204.

[2] Pati, M. Parhi, M. Alnabhan, B. K. Pattanayak, A. K. Habboush, and M. K. Al Nawayseh, “An IoT-Fog-Cloud Integrated Framework for Real-Time Remote Cardiovascular Disease Diagnosis,” Informatics, vol. 10, no. 1, p. 21, Feb. 2023, doi: 10.3390/informatics10010021.

[3] Pati, M. Parhi and B. K. Pattanayak, "IDMS: An Integrated Decision Making System for Heart Disease Prediction," 2021 1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON), Bhubaneswar, India, 2021, pp. 1-6, doi: 10.1109/ODICON50556.2021.9428958.

[4] Pati, M. Parhi, and B. K. Pattanayak, “HeartFog: Fog Computing Enabled Ensemble Deep Learning Framework for Automatic Heart Disease Diagnosis,” Smart Innovation, Systems and Technologies, pp. 39–53, 2022, doi: https://doi.org/10.1007/978-981-16-9873-6_4.

[5] A. Pati, M. Parhi, and B. K. Pattanayak, “IHDPM: an integrated heart disease prediction model for heart disease prediction,” International Journal of Medical Engineering and Informatics, vol. 14, no. 6, p. 564, 2022, doi: https://doi.org/10.1504/ijmei.2022.126526.

[6] D. S. K. Nayak, A. Pati, A. Panigrahi, S. Sahoo, and T. Swarnkar, “ReCuRandom: A hybrid machine learning model for significant gene identification,” Nucleation and Atmospheric Aerosols, Jan. 2023, doi: https://doi.org/10.1063/5.0137029.

[7] B. Sahu, A. Panigrahi, and S. K. Rout, “DCNN-SVM: A New Approach for Lung Cancer Detection,” CRC Press eBooks, Jun. 2020, doi: https://doi.org/10.1201/9781003043980-13.

[8] P. Chatterjee, L. J. Cymberknop and R. L. Armentano, "IoT-based decision support system for intelligent healthcare — applied to cardiovascular diseases," 2017 7th International Conference on Communication Systems and Network Technologies (CSNT), Nagpur, India, 2017, pp. 362-366, doi: 10.1109/CSNT.2017.8418567.

[9] F. Jabeen et al., “An IoT based efficient hybrid recommender system for cardiovascular disease,” Peer-to-Peer Networking and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 10
DOI: https://doi.org/10.17762/ijritcc.v11i10.8483
***Article Received:** 29 July 2023 **Revised:** 19 September 2023 **Accepted:** 01 October 2023*

Applications, vol. 12, no. 5, pp. 1263–1276, Mar. 2019, doi: https://doi.org/10.1007/s12083-019-00733-3.

[10] T. Nguyen Gia et al., “Energy efficient fog-assisted IoT system for monitoring diabetic patients with cardiovascular disease,” Future Generation Computer Systems, vol. 93, pp. 198–211, Apr. 2019, doi: https://doi.org/10.1016/j.future.2018.10.029.

[11] S. P. Patro, N. Padhy, and D. Chiranjevi, “Ambient assisted living predictive model for cardiovascular disease prediction using supervised learning,” Evolutionary Intelligence, Sep. 2020, doi: https://doi.org/10.1007/s12065-020-00484-8.

[12] M. Ahmid and O. Kazar, "A Cloud-IoT Health Monitoring System Based on Smart Agent for Cardiovascular Patients," 2021 International Conference on Information Technology (ICIT), Amman, Jordan, 2021, pp. 1-6, doi: 10.1109/ICIT52682.2021.9491113.

[13] C. Huang, Y. Zong, J. Chen, W. Liu, J. Lloret and M. Mukherjee, "A Deep Segmentation Network of Stent Structs Based on IoT for Interventional Cardiovascular Diagnosis," in IEEE Wireless Communications, vol. 28, no. 3, pp. 36-43, June 2021, doi: 10.1109/MWC.001.2000407.

[14] P. Wang et al., “A Wearable ECG Monitor for Deep Learning Based Real-Time Cardiovascular Disease Detection,” arXiv (Cornell University), Jan. 2022, doi: https://doi.org/10.48550/arxiv.2201.10083.

[15] “UCI Machine Learning Repository: Heart Disease Data Set.” Available: http://archive.ics.uci.edu/ml/datasets/Heart+Disease (accessed Mar. 23, 2023).

[16] D. Mohapatra, J. Tripathy, K. K. Mohanty and D. S. K. Nayak, "Interpretation of Optimized Hyper Parameters in Associative Rule Learning using Eclat and Apriori," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2021, pp. 879-882, doi: 10.1109/ICCMC51019.2021.9418049.

[17] T. A. Retson, A. H. Besser, S. Sall, D. Golden, and A. Hsiao, “Machine Learning and Deep Neural Networks in Thoracic and Cardiovascular Imaging,” Journal of Thoracic Imaging, vol. 34, no. 3, pp. 192–201, May 2019, doi: https://doi.org/10.1097/rti.0000000000000385.

[18] B. Sahu, A. Panigrahi, S. Mohanty, and S. S. Panigrahi, “A hybrid Cancer Classification Based on SVM Optimized by PSO and Reverse Firefly Algorithm,” International Journal of Control and Automation, vol. 13, no. 4, pp. 506–517, May 2020.

[19] A. Pati, M. Parhi, and B. K. Pattanayak, “COVID-19 Pandemic Analysis and Prediction Using Machine Learning Approaches in India,” Lecture Notes in Networks and Systems, pp. 307–316, 2021, doi: https://doi.org/10.1007/978-981-16-0695-3_30.

[20] B. Sahu, A. Panigrahi, S. Pani, S. Swagatika, D. Singh and S. Kumar, "A Crow Particle Swarm Optimization Algorithm with Deep Neural Network (CPSO-DNN) for High Dimensional Data Analysis," 2020 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2020, pp. 0357-0362, doi: 10.1109/ICCSP48568.2020.9182181.

[21] A. Pati, M. Parhi, and B. K. Pattanayak, “A review on prediction of diabetes using machine learning and data mining classification techniques,” International Journal of Biomedical Engineering and Technology, vol. 41, no. 1, p. 83, 2023, doi: https://doi.org/10.1504/ijbet.2023.128514.

[22] A. Pati, M. Parhi, and B. K. Pattanayak, “ensemble deep learning approach for Chronic kidney disease (CKD) prediction,” Nucleation and Atmospheric Aerosols, Jan. 2023, doi: https://doi.org/10.1063/5.0136894.

[23] J. Tripathy, R. Dash, B. K. Pattanayak, S. K. Mishra, T. K. Mishra, and D. Puthal, “Combination of Reduction Detection Using TOPSIS for Gene Expression Data Analysis,” Big Data and Cognitive Computing, vol. 6, no. 1, p. 24, Feb. 2022, doi: 10.3390/bdcc6010024.

[24] A. Pati, M. Parhi, and B. K. Pattanayak, “An ensemble approach to predict acute appendicitis," 2022 International Conference on Machine Learning, Computer Systems and Security (MLCSS), Bhubaneswar, India, 2022, pp. 183-188, doi: 10.1109/MLCSS57186.2022.00041.

[25] B. Sahu, A. Panigrahi, B. Dash, P. K. Sharma, and A. Pati, “A hybrid wrapper spider monkey optimization-simulated annealing model for optimal feature selection,” International Journal of Reconfigurable and Embedded Systems (IJRES), vol. 12, no. 3, pp. 360–375, Nov. 2023, doi: https://doi.org/10.11591/ijres.v12.i3.pp360-375.

[26] A. Pati, A. Panigrahi, D. S. K. Nayak, G. Sahoo, and D. Singh, “Predicting Pediatric Appendicitis using Ensemble Learning Techniques,” Procedia Computer Science, vol. 218, pp. 1166–1175, 2023, doi: https://doi.org/10.1016/j.procs.2023.01.095.

[27] A. Panigrahi et al., "En-MinWhale: An ensemble approach based on MRMR and Whale optimization for Cancer diagnosis," in IEEE Access, doi: 10.1109/ACCESS.2023.3318261.

[28] A. Pati et al., “FOHC: Firefly Optimizer Enabled Hybrid approach for Cancer Classification,” International Journal on Recent and Innovation Trends in Computing and Communication, vol. 11, no. 7s, pp. 118–125, Jul. 2023, doi: https://doi.org/10.17762/ijritcc.v11i7s.6983.

[29] S. S. Tripathy, N. Tripathy, M. Rath, A. Pati, A. Panigrahi and M. Dash, "Smart Farming based on Deep Learning Approaches," 2022 2nd Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON), Bhubaneswar, India, 2022, pp. 1-5, doi: 10.1109/ODICON54453.2022.10010165.

[30] S. K. Rout, B. Sahu, A. Panigrahi, B. Nayak, and A. Pati, “Early Detection of Sepsis Using LSTM Neural Network with Electronic Health Record,” Smart Innovation, Systems and Technologies, pp. 201–207, Nov. 2022, doi: https://doi.org/10.1007/978-981-19-6068-0_19.

[31] S. Joshi, I. S. Samanta, S. Panda, P. K. Rout and K. Swain, "Breast Cancer Detection Using Variational Mode Decomposition (VMD) and Weighted Bidirectional Extreme Learning Machine," 2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT), Bhubaneswar, India, 2023, pp. 369-372, doi: 10.1109/APSIT58554.2023.10201695.